



SPOLM 2009

ISSN 2175-6295

Rio de Janeiro- Brasil, 05 e 06 de agosto de 2009.

043/2009 - A UTILIZAÇÃO DE MÉTODOS ESTATÍSTICOS MULTIVARIADOS COMO FERRAMENTA DE ANÁLISE ECONÔMICO-FINANCEIRA DE EMPRESAS PARA CONCESSÃO DE CRÉDITO

Amilton Fernando Cardoso

ITA - Instituto Tecnológico de Aeronáutica
Pça. Mal. Eduardo Gomes, 50; 12.228-900 – São José dos Campos/SP.
amilthon@terra.com.br

Arnoldo Souza Cabral

ITA - Instituto Tecnológico de Aeronáutica
Pça. Mal. Eduardo Gomes, 50; 12.228-900 – São José dos Campos/SP.
cabral@ita.br

Resumo

O objetivo do artigo é analisar o desempenho de empresas para fins de concessão de crédito, por meio das variáveis contábeis. A metodologia do estudo propôs uma análise estatística dos dados (teste T^2 de Hotelling e descartes de *outliers* e variáveis), seguida dos métodos discriminantes: Função Discriminante Linear de Fisher (FDLF) e Modelo da Regressão Logístico Múltiplo (MRLM). Na avaliação dos modelos foi utilizado o método de Lachenbruch. A implementação de funções no software MATLAB® possibilitou a determinação dos resultados necessários. Consideram-se como resultados importantes encontrados: os descartes terem sido de fundamental importância e, principalmente, a eficiência dos modelos para o objetivo proposto, pois resultados promissores foram encontrados MRLM (93,18%) e FDLF (91,67%).

Palavras-chave: reconhecimento de padrões; análise de crédito; análise multivariada.

Abstract

The objective of the article is to analyze the performance of companies of credit concession, through accounting variables. The methodology of the study proposes, firstly, a statistical analysis of the data (T^2 test of Hotelling and discarded outliers and variables) and then discriminatory methods: Fisher Discriminating Linear Function (FDLF) and Lachenbruch Multiple Logistic Regression Model. An implementation of functions in MATLAB® software provided the establishment of the needed results. It is considered as important results found: the discards having been of fundamental importance and, most of all, the efficiency of models for the pre-set aim. As promising results it was found: MRLM (93.18%) and FDLF (91.67%).

Keywords: pattern precognition; credit analysis; multivariate analysis.

1. INTRODUÇÃO

No atual contexto da economia brasileira figuram os juros altos, o aumento do desemprego e uma pesada carga tributária. Todos esses fatores concorrem para um crescimento mais modesto da economia brasileira e para a redução da oferta de crédito. A falta de crédito é um dos principais obstáculos para a geração de novos empregos. Por outro lado, a existência de crédito a uma taxa de juros aceitável proporcionaria mais empregos e melhores condições de crescimento econômico ao país. Dessa forma, percebe-se a grande responsabilidade social em torno da questão.

A inadimplência é o principal fator que aflige qualquer instituição financeira. Afinal, a aprovação dos créditos e a definição das taxas a serem cobradas são decorrentes dos riscos que envolvem a carteira de operações (*portfólios*).

Para Steiner *et al.* (1999) [1], qualquer erro na decisão de concessão de crédito pode significar que em uma única operação haja a perda do ganho obtido em dezenas de outras bem-sucedidas, visto que o não recebimento representa a perda do montante emprestado.

Assim, é importante prever a inadimplência, pois os prejuízos com operações mal sucedidas provavelmente serão cobertos com a cobrança de uma sobretaxa a novos créditos, ou seja, encarecem-se as operações futuras. Matarazzo (1998) [2] cita que é comum dois analistas de balanços chegarem a conclusões diferentes a respeito de balanços de uma mesma empresa.

Logo, um processo decisório, sem se basear no *feeling* do tomador de decisões, torna-se de fundamental importância.

2. METODOLOGIA DA PESQUISA

Por meio de índices econômico-financeiros coletados dos demonstrativos contábeis de empresas – Balanço Patrimonial (BP) e Demonstração do Resultado do Exercício (DRE) –, classificá-las como adimplentes ou inadimplentes, visando diminuir o risco de inadimplência.

São propostas técnicas estatísticas multivariadas, a citar: Função Discriminante Linear de Fisher (FDLF), Modelo de Regressão Logístico Múltiplo (MRLM) e, ainda, um método de avaliação desses procedimentos (abordagem de Lachenbruch – metodologia *jackknife*).

Consideram-se também a validade do tratamento estatístico dos dados, preliminar à aplicação dos métodos discriminatórios propostos; e a construção de um programa computacional, no software MATLAB®, para determinar os resultados dos testes estatísticos e estimar os parâmetros dos modelos discriminatórios propostos, além da avaliação de precisão desses modelos construídos.

Assumiu-se, como adimplência ou inadimplência, a capacidade ou incapacidade de a empresa pagar as obrigações financeiras quando elas vencem. Parte-se do pressuposto da veracidade das informações contidas nos indicadores contábeis.

A partir da seleção de alguns índices considerados relevantes pela literatura, utilizam-se escores das duas primeiras componentes principais padronizadas para o descarte de observações (*outlier*), caso existam.

Os índices selecionados serão submetidos à eventual necessidade de redução do número de variáveis (via componentes principais), ou seja, consideram-se diversas variáveis e reduz-se, caso necessário, o conjunto de variáveis para um subconjunto que possua maior precisão, por meio dos autovalores e autovetores da matriz de correlação.

O teste de hipótese, T2 de Hotelling, é aplicado com a finalidade de verificar se as duas populações representadas por suas amostras são provenientes de populações distintas, ou seja, se existe diferença nas suas várias características médias.

2.1. AMOSTRA DA PESQUISA

A amostra caracteriza-se por empresas saudáveis e empresas com problemas. Os dados de empresas, utilizados neste trabalho, foram selecionados dos arquivos de uma agência de fomento, situada na região sul do Brasil, de forma aleatória, contendo informações de empresas solventes e empresas insolventes.

A amostra foi tomada de maneira sistemática, com intervalo padronizado por um auditor do banco, resguardando-se os nomes das empresas, por exigência do próprio banco.

Dadas às características das empresas analisadas, e sendo as mesmas potenciais tomadoras de recurso junto ao banco, há uma grande concentração de empresas consideradas saudáveis (de um total de 136 empresas, escolhidas aleatoriamente, há 118 empresas adimplentes e apenas 18 empresas inadimplentes).

2.2 INDICADORES ECONÔMICO-FINANCEIROS

O quadro 1 identifica os indicadores financeiros e econômicos e as 8 variáveis utilizadas inicialmente para o tratamento estatístico dos dados.

Índices de Estrutura de Capitais (Retirados do Balanço Patrimonial (BP) - Ind. financeiros)
Participação de Capital de Terceiros ($PCT = PC + ELP / PL$)
Imobilização do Patrimônio Líquido ($IPL = AP / PL$)
Índices de Liquidez (Retirados do Balanço Patrimonial - Ind. financeiros)
Liquidez Seca ($LS = AC - Est - Des / PC$)
Liquidez Corrente ($LC = AC / PC$)
Liquidez Geral ($LG = AC + RLP / PC + ELP$)
Margem de Garantia ($MG = AT / PC + ELP$)
Índices de Rentabilidade (Retirados da Demonstração do Resultado do Exercício (DRE) - Ind. Econômicos)
Retorno sobre Vendas ($RsV = LL / VL$)
Retorno sobre Patrimônio Líquido ($RPL = LL / PL - LL$)
Sendo:
AC = Ativo Circulante; PC = Passivo Circulante; AP = Ativo Permanente; AT = Ativo Total
ELP = Exigível a Longo Prazo; RLP = Realizável a Longo Prazo; VL = Vendas Líquidas
PL = Patrimônio Líquido; LL = Lucro Líquido; Des = Despesas do Exercício Seguinte;
Est = Estoques; Ind = Indicadores

Fonte: Adaptado de Lima (2002) [3]

Quadro 1 - Indicadores financeiro-econômicos

Tomando como ponto de partida os dados das 136 empresas contendo os valores para as variáveis apresentadas anteriormente e aplicando o descarte de outlier (via componentes principais), verificou-se que 4 empresas possuíam valores discrepantes (ponto de vista contábil) em alguns índices, sendo as mesmas descartadas do banco de dados.

De posse das 132 empresas restantes, aplicou-se o descarte de variáveis (também se utilizando de componentes principais) para verificar a necessidade de redução do número de variáveis, tendo como resultado a possibilidade de se trabalhar apenas com as seguintes variáveis: MG, IPL, LS e RPL.

Logo, a matriz inicial, de ordem 136x8, foi reduzida a uma matriz cujas dimensões passaram a ser 132x4, reduzindo significativamente o banco de dados. A aplicação da Função Discriminante Linear de Fisher e do Modelo de Regressão Logístico Múltiplo se deu a esse banco de dados.

Portanto, os métodos de análise discriminante multivariada foram aplicados a duas populações munidas de uma partição definida a priori, descritas por 4 variáveis explicativas (ou preditoras), e o objetivo principal foi construir uma regra de decisão utilizando uma amostra constituída de 132 empresas descritas pelas 4 variáveis, sendo conhecido o grupo a priori a que pertenciam.

2.3. FUNÇÃO DISCRIMINANTE LINEAR DE FISHER (FDLF)

Para Johnson e Wichern (1998) [4], a FDLF, basicamente, consiste em separar duas classes de objetos ou fixar um novo objeto em uma das duas classes. Ao definir a FDLF, é comum denominar as classes (populações) de π_1 e π_2 , e os objetos, separados ou classificados com base nas medidas de p variáveis aleatórias são associados com vetores do tipo $X' = [X_1, X_2, \dots, X_p]$.

O objetivo de Fisher, ao criar essa regra de reconhecimento de padrões e classificação, foi transformar as observações multivariadas X's nas observações univariadas Y's, tal que os Y's das populações π_1 e π_2 fossem separados em relação às médias das duas populações tanto quanto possível. Fisher teve a idéia de tomar combinações lineares de X para criar os Y's, dado que essas combinações são funções de X e, por outro lado, são facilmente calculáveis (ZUGE e CHAVES NETO, 1999) [5].

Seja μ_{1Y} a média dos Y's obtidos dos X's pertencentes a π_1 e μ_{2Y} a média dos Y's obtidos dos X's pertencentes a π_2 e, considerando a matriz de covariância, Σ , como sendo a mesma para ambas as populações, então Fisher selecionou a combinação linear que maximiza a razão entre o "quadrado da distância entre as médias" e a "variância de Y", ou seja:

$$\frac{(\mu_{1Y} - \mu_{2Y})^2}{\sigma_Y^2} = \frac{[C'(\mu_1 - \mu_2)]^2}{C' \Sigma C}$$

Essa razão é maximizada por: $C = \Sigma^{-1}(\mu_1 - \mu_2)$ e, assim, tem-se: $Y = C'X = (\mu_1 - \mu_2)' \Sigma^{-1} X$, que é conhecida como FDLF populacional.

Como, na realidade, os parâmetros μ_1, μ_2 e Σ não são conhecidos, devem-se usar seus estimadores, ou seja, \bar{X}_1 estima μ_1 , \bar{X}_2 estima μ_2 e a matriz de covariância conjunta (estimada) é dada por:

$$S_p = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2}$$

Dessa forma, a FDLF amostral fica determinada por:

$$Y = (\bar{X}_1 - \bar{X}_2)' S_p^{-1} X$$

A estimativa do ponto médio amostral é dada por:

$$\hat{m} = \frac{1}{2} [(\bar{X}_1 - \bar{X}_2)' S_p^{-1} (\bar{X}_1 + \bar{X}_2)]$$

Obtém-se, desta forma, a seguinte regra de classificação:

- Alocar X_0 em π_1 se $\hat{y}_0 = (\bar{X}_1 - \bar{X}_2)' S_p^{-1} X_0 \geq \hat{m}$
ou
- Alocar X_0 em π_2 ,
se $\hat{y}_0 = (\bar{X}_1 - \bar{X}_2)' S_p^{-1} X_0 < \hat{m}$

2.4. MODELO DE REGRESSÃO LOGÍSTICO MÚLTIPLO (MRLM)

O Modelo de Regressão Logístico Múltiplo é uma técnica estatística comumente usada para a análise de dados com resposta binária (dicotômica) e consiste em relacionar, por meio de um modelo, a variável resposta com os fatores que influenciam a ocorrência de um determinado evento. Por exemplo, o quanto os índices de estrutura de capitais, liquidez e rentabilidade explicam o risco de inadimplência.

Como no presente estudo a variável resposta é dicotômica, o modelo de regressão linear não deve ser utilizado, pois poder-se-á obter valores de resposta estimada fora do intervalo $[0,1]$, não sendo neste caso compatível com a natureza do fenômeno estocástico.

Além disso, facilmente se mostra que a condição de variância constante para os resíduos não se verifica. Sejam uma variável aleatória Y (dicotômica, ou seja, assumindo apenas os valores 0 ou 1) e $X' = [X_1, X_2, \dots, X_p]$ um vetor de dimensão p , composto de variáveis aleatórias independentes, e ainda tomando-se n observações independentes, pode-se escrever o Modelo de Regressão Logística Múltiplo na forma:

$$P(X) = \frac{e^{\beta'X}}{1 + e^{\beta'X}}$$

para

$$\beta' = [\beta_0, \beta_1, \dots, \beta_p] \text{ e } X = [1, X_1, \dots, X_p]$$

Este modelo é o mais apropriado no caso de a variável dependente ser dicotômica, pois a Regressão Logística estima diretamente a probabilidade de ocorrência de um evento.

A estimação dos parâmetros deve ser feita pelo método de máxima verossimilhança, sendo o mais recomendado quando se dispõem de observações individuais da ocorrência ou não de determinado evento. Esse método fornece estimativas para os parâmetros que maximizam a probabilidade de obter o conjunto observado de dados.

Assim, tomando uma amostra aleatória composta de n pares de observações (X_i, Y_i) com $i = 1, 2, \dots, n$, onde os Y 's representam os valores observados de uma variável dicotômica e os X 's representam os valores observados de p variáveis independentes.

Uma forma conveniente para expressar a contribuição da função de verossimilhança para os pares (X_i, Y_i) é por meio da fórmula:

$$F(X_i) = P(X_i)^{Y_i} [1 - P(X_i)]^{1 - Y_i}$$

Desde que assumido que as observações são independentes, a função de verossimilhança é obtida como o produto dos termos dados na equação anterior:

$$l(\beta) = \prod_{i=1}^n F(X_i) = \prod_{i=1}^n [P(X_i)]^{Y_i} [1 - P(X_i)]^{1 - Y_i}$$

Assim, a função de log-verossimilhança fica determinada por:

$$L(\beta) = \sum_{i=1}^n [Y_i \beta' X_i - \ln(1 + e^{\beta' X_i})]$$

Desta forma, estimar-se-ão (p+1) equações de verossimilhança, as quais são obtidas por derivadas da função de log-verossimilhança com respeito à (p+1) coeficientes. As equações de verossimilhança resultantes podem ser representadas como:

$$\frac{\partial L(\beta)}{\partial \beta_0} = \sum_{i=1}^n [Y_i - P(X_i)] = 0$$

e

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^n X_{ij} [Y_i - P(X_i)] = 0$$

Para $j = 1, 2, \dots, p$, o que reflete no fato de a soma dos valores observados de Y ser igual à soma dos valores esperados (preditos), mediante a aplicação do modelo obtido. Esta propriedade é importante, pois ajuda a avaliar o ajuste do modelo.

A estimação dos parâmetros para o MRLM foi feita com base nos estimadores de Máxima Verossimilhança, e o seu cálculo foi obtido por meio do método de Levenberg-Marquardt. Em termos da aferição da qualidade global do modelo logístico, foi utilizado o critério baseado nos resultados de classificação obtidos com esse modelo.

O método de Levenberg-Marquardt foi utilizado para determinar os estimadores de máxima verossimilhança do modelo de regressão logístico múltiplo. Esse método trabalha com uma variação suave entre os extremos dos métodos de Newton e do Gradiente, sendo que se utiliza este último quando se está longe do mínimo, visto que o método de Newton pode não convergir nesta situação.

Por outro lado, quando se está distante da solução usa-se o método do Gradiente. Esse método combina os métodos do Gradiente e de Newton por meio da alteração dos elementos da diagonal principal da matriz *hessiana*.

No presente estudo, o critério de parada foi estabelecido a partir de duas considerações: a primeira se relaciona ao fato de que, analisando as derivadas parciais da função de log-verossimilhança, tem-se que a soma dos valores observados deve ser igual à soma dos valores preditos, e a segunda está baseada no fato de uma nova solução não trazer alteração no valor da função objetivo (log-verossimilhança). Assim, o processo iterativo é suspenso apenas quando as duas condições descritas anteriormente forem satisfeitas.

2.5. AVALIAÇÃO DA FUNÇÃO DE CLASSIFICAÇÃO: A ABORDAGEM DE LACHENBRUCH

O método de Lachenbruch, Johnson e Wichern (1998) [4] mencionam que é uma abordagem estatística empregada para avaliar o desempenho da regra de reconhecimento de padrões e estimar a probabilidade de classificação correta e incorreta em cada grupo, ou seja, uma forma de avaliar a eficiência da regra de classificação.

O algoritmo do referido procedimento é descrito a seguir:

- a) Passo 1: Inicie com as observações do grupo π_1 , omita uma observação desse grupo e construa uma função de classificação usando as (n_1-1) restantes do grupo π_1 acrescidas das n_2 observações do grupo π_2 .
- b) Passo 2: Classifique a observação retida, usando a função construída no passo 1.

c) Passo 3: Repita os passos 1 e 2 até que todas as observações de π_1 sejam classificadas. Seja $n_{2/1}$ o número de observações reconhecidas erroneamente nesse grupo.

d) Passo 4: Repita os passos de 1 a 3 para as n_2 observações do grupo π_2 . Seja $n_{1/2}$ o número de observações reconhecidas erroneamente nesse grupo.

- $\hat{P}(2/1) = \frac{n_{2/1}}{n_1}$ (classificar em π_2 , quando a observação pertence a π_1)
- $\hat{P}(1/2) = \frac{n_{1/2}}{n_2}$ (classificar em π_1 , quando a observação pertence a π_2)

Desta forma, a proporção esperada de erro é dada por:

$$\hat{E}(\text{AER}) = \frac{n_{1/2} + n_{2/1}}{n_1 + n_2}$$

3. ANÁLISE DOS RESULTADOS

Na tabela 1 são especificados os principais resultados obtidos para a análise da Função Discriminante Linear de Fisher.

VARIÁVEL	COEFICIENTE DA FDLF (C)
MG = AT / (PC + ELP)	0.0828
IPL = AP / PL	-0.0235
LS = (AC – Est – Des) / PC	0.3004
RPL = LL / PL	0.0522

Fonte: Adaptado de Lima (2002) [3]

Tabela 1 – Coeficientes da FDLF

Assim, a aplicação da metodologia FDLF, na amostra em estudo, resultou na seguinte função de classificação: $Y=0.0828*MG - 0.0235*IPL+0.3004*LS+0.0522*RPL$

Desta forma, ao multiplicar esse vetor C' de dimensão quatro (4 indicadores) pelos valores de uma nova observação, tem-se um escore. Se esse for maior que o ponto médio (\hat{m}), classificar-se-á na população das empresas adimplentes; caso contrário deve ser classificado no grupo das empresas inadimplentes. O ponto médio a ser utilizado para fazer a decisão de classificação foi estimado por: $\hat{m} = -3.7946$.

A probabilidade de classificação correta para a FDLF é de 91,67%, avaliada pelo método de Lachenbruch. A tabela 2 demonstra os resultados de classificação para A FDLF

POPULAÇÃO	Nº DE EMPRESAS	TAXA (%)
População 1		
Certo	109	94,78
Errado	6	5,22
População 2		
Certo	12	70,59
Errado	5	29,41

Probabilidade de classificação correta = 91.67%

Fonte: Adaptado de Lima (2002) [3]

Tabela 2 – Resultados de classificação para A FDLF

O resultado encontrado para a FDLF é melhor do que o resultado que inclui as 132 empresas com suas 8 variáveis, ou seja, sem efetuar o descarte de variáveis.

Assim, considera-se de grande utilidade o descarte de variáveis, desde que bem fundamentado, pois neste estudo o número de medidas (variáveis) foi reduzido em 50% e resultados promissores foram encontrados.

A tabela 3 apresenta os coeficientes e os erros padrões estimados para o Modelo de Regressão Logístico Múltiplo, tendo por base os estimadores de máxima verossimilhança obtidos por meio do método de Levenberg-Marquardt.

VARIÁVEL	COEFICIENTE DO MRLM	ERRO PADRÃO
Constante	0,0789	1,2686
MG	0,4717	0,5385
IPL	-0,0133	0,0047
LS	5,1297	1,8334
RPL	0,0465	0,0188

Fonte: Adaptado de Lima (2002) [3]

Tabela 3 – Coeficientes do MRLM

Assim a aplicação do MRLM, na amostra em estudo, resultou no seguinte modelo:

$$P(X) = \frac{e^{0.0789 + 0.4717 * MG - 0.0133 * IPL + 5.1297 * LS + 0.0465 * RPL}}{1 + e^{0.0789 + 0.4717 * MG - 0.0133 * IPL + 5.1297 * LS + 0.0465 * RPL}}$$

O valor obtido por P(X) representa a probabilidade de a empresa ser adimplente. O valor *cut off score*, o qual maximiza a probabilidade de acerto, foi igual a 0.5. A tabela 4 apresenta os resultados de classificação para o MRLM, avaliado pela abordagem de Lachenbruch.

POPULAÇÃO	N.º DE EMPRESAS	TAXA (%)
POPULAÇÃO 1		
Certo	110	95,65
Errado	5	4,35
POPULAÇÃO 2		
Certo	13	76,47
Errado	4	23,53

NOTA: Probabilidade de classificação correta 93,18= %

Fonte: Adaptado de Lima (2002) [3]

Tabela 4 - Resultados de classificação para o MRLM

De acordo com as tabelas 2 e 4, observa-se uma porcentagem de classificação correta superior, por meio do Modelo de Regressão Logístico Múltiplo, em todos os quesitos avaliados.

A taxa de erro tipo II teve uma redução significativa, pois esse tipo de relevante, pois está se concedendo empréstimo a uma empresa que não vai pagar, ocasionando a perda dos juros e do montante emprestado.

Um resultado comparativo entre os métodos propostos (FDLF e MRLM), obtido por meio da abordagem de Lachenbruch. O quadro 2 destaca a comparação entre os modelos desenvolvidos

POPULAÇÃO / MÉTODO	FDLF		MRLM	
	Abs.	%	Abs.	%
Pop. (1 1)	109	94,78	110	95,65
Pop. (2 1)	6	5,22	5	4,35
Pop. (2 2)	12	70,59	13	76,47
Pop. (1 2)	5	29,41	4	23,53
TOTAL	132	100,00	132	100,00
Acerto	121/132 =	91,67	123/132 =	93,18

Fonte: Adaptado de Lima (2002) [3]

Quadro 2 - Comparação entre os modelos desenvolvidos

A notação POP(i|j) representa o número de elementos da população j que foi classificado na população i, estando os mesmos corretamente classificados quando $i = j$, e incorretamente, caso contrário.

4. CONCLUSÃO

Os resultados positivos encontrados com a estimação dos parâmetros após a análise preliminar dos dados (T^2 de Hotelling e os descartes de outlier e variáveis). Portanto, é relevante uma análise estatística rigorosa antes da aplicação dos métodos de discriminação.

O MRLM foi o que apresentou melhor desempenho, quando analisado em relação ao percentual de classificação correta em cada um dos itens avaliados (erros tipo I e tipo II, e acerto global). Por outro lado, a FDLF apresentou um percentual de classificação correta

inferior ao MRLM, mas esta técnica é bastante eficiente, pois sua percentagem de classificação correta foi superior a 91%.

Conclui-se que a FDLF e o MRLM são duas técnicas eficientes na discriminação das empresas adimplentes e inadimplentes, evidenciando certa melhora para a segunda técnica em relação à primeira. Apesar desta evidência, não se pode contestar a eficiência das referidas técnicas para o objetivo proposto.

A diferença verificada em relação à eficiência dos dois procedimentos desenvolvidos não constitui empecilho à sua utilização. Ao contrário, possibilita ao tomador de decisões a escolha da técnica que apresente uma maior probabilidade de classificação correta, o que fortalece o fator segurança, tendo em vista o montante de recursos envolvidos nas operações de créditos às empresas.

As técnicas abordadas neste trabalho, se empregadas corretamente, constituem-se em eficientes ferramentas auxiliares dos gestores de crédito, ou seja, os modelos desenvolvidos poderão auxiliar os analistas de créditos bancários no processo de tomada de decisões e previsão do comportamento de futuras empresas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

[4] JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. 4. ed. Nova Jersey: Prentice-Hall, inc., 1998.

[3] LIMA, J. D. de. A análise econômico-financeira de empresas sob a ótica da estatística multivariada. Curitiba, 2002. 192 f. *Dissertação (Mestrado em Métodos Numéricos em Engenharia)* – Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná.

[2] MATARAZZO, D. C. *Análise financeira de balanços: abordagem básica e gerencial*. 5. ed., São Paulo: Atlas, 1998.

[1] STEINER, M. T. A.; CARNIERI, C.; STEINER NETO, P. J.; KOPITTKE, B. *Sistemas especialistas probabilístico e redes neurais na análise do crédito bancário*. Revista de Administração, São Paulo, v. 34, n. 3, p. 56-67, jul./set. 1999.

[5] ZÜGE, M.; CHAVES NETO, A. *Utilização dos métodos estatísticos multivariados na avaliação do desempenho empresarial*. Revista Paranaense de Desenvolvimento, Economia e Sociedade, Curitiba, n. 97, p. 101-112, set./dez. 1999.